

Evaluating Semi-automatic Annotation of Domestic Energy Consumption as a Memory Aid

Darren P. Richardson, Enrico Costanza, Sarvapali D. Ramchurn

University of Southampton, United Kingdom
dpr1g09@zepler.net, {ec, sdr}@ecs.soton.ac.uk

ABSTRACT

Frequent feedback about energy consumption can help conservation, one of the current global challenges. Such feedback is most helpful if users can relate it to their own day-to-day activities. In earlier work we showed that manual annotation of domestic energy consumption logs aids users to make such connection and discover patterns they were not aware of. In this poster we report how we augmented manual annotation with machine learning classification techniques. We propose the design of a lab study to evaluate the system, extending methods used to evaluate context aware memory aids, and we present the results of a pilot with 5 participants.

Author Keywords Energy management; eco-feedback; sustainable HCI; consumption, pattern recognition; human-agent interaction; memory aid.

ACM Classification Keywords H.5.2 Information Interfaces and Presentation: User Interfaces

General Terms Experimentation, Human Factors

INTRODUCTION

We are interested in helping users understand their own energy consumption, with the long term goal of encouraging conservation. Indeed prior work indicates that providing users with frequent feedback about their energy consumption can be helpful for conservation, but there is a need to help users relate the feedback to their day-to-day activities [3]. In prior work [1] we showed that interactively annotating energy consumption logs can help users to relate consumption to everyday activities, and to discover patterns they were not aware of. Using a combination of low-cost, off-the-shelf networked electricity sensors and web technologies, the FigureEnergy system enables users to annotate their own energy logs and integrates these annotations in multiple interactive visualizations. FigureEnergy runs in a standard web browser and annotation takes place simply by dragging the mouse over an interval of the energy logs and associating an icon and a free text label to this interval. We refer to these annotated intervals as *events*. Annotated events are displayed in terms of their total energy consumption, making it easier to compare them.

One limitation of manual annotation is that it requires a

considerable amount of a user's time and effort, with a risk of reducing the user's motivation to use the system in the long term. As an attempt to overcome these problems and reduce the amount of effort required from users, in this poster we report the integration of pattern recognition techniques to assist the energy logs annotation and we propose an evaluation strategy for the system.

INTRODUCING SEMI-AUTOMATIC ANNOTATIONS

As for the initial FigureEnergy project, the purpose of this semi-automatic extension is to draw a user's attention to their electricity usage patterns, and help them understand how their different day-to-day activities impact on their overall energy consumption. Zeifman and Roth [5] have reviewed the field of Non-intrusive Appliance Load Monitoring (NIALM), and shown that whilst there are many issues still to be solved, machine learning techniques can be used on energy consumption data to distinguish different classes of appliance, to a certain level of accuracy. Therefore, we decided to implement a simpler system using an "off-the-shelf" open-source machine learning library to perform energy usage classification. Even though similar to NIALM, the problem we address is distinct: rather than attempting to disaggregate the total consumption *objectively* into separate appliances, our aim is to assist users in the task of *subjective* annotation of their consumption. Initially, potential events are *detected* in the energy consumption signal by selecting all periods where the power was greater than a threshold (the mean plus one standard deviation), and then merging events occurring within 4 minutes of one another to account for multi-state appliances such as dishwashers and washing machines. We then attempt to classify these potential events through a support vector machine trained on each user's annotations, using the following features of the events: duration; mean power consumption; maximum power consumption; minimum power consumption; standard deviation of power consumption; time of day; and day of week. In this way, it is possible to predict which label the user would be most likely to assign to each automatically detected event, and offer this as a suggestion to the user, with an aim to make the annotation task easier and more accurate.

The system was trained and tested on the data collected during the FigureEnergy field trial from 6 users who created at least 20 annotations. The classifier yielded a mean accuracy of 59.3% (s.d. 18.1%, ranging from 34.6% to 88.9% per user). The overall average precision of the detection system was 42.0% and the recall was 72.4%. The

low precision is not surprising, as the users from the field trial rarely annotated all of the usage events.

When integrating this with the user interface, we decided that we wanted the classifications to be subtle: clearly visible, but not intrusive and certainly not making any decisions that should be taken by the user. The design rationale being that if the automation is too “automatic,” then users' familiarity with their electricity consumption may fall dramatically, rendering the system pointless. We therefore decided to have the predictions displayed as suggestions that the user could take into account or ignore. This way the user still has to make the final decision for each event, keeping them involved with the annotation process. From this perspective the suggestions are designed to act as memory aids for the user, to remind them which activities or appliances contributed to the consumption, if they had forgotten.

STUDY DESIGN

Background

An earlier field trial [1] demonstrated that the interactive annotation of energy consumption offered by FigureEnergy is successful overall, so we want to determine whether augmenting the system with automatic suggestions could help users or get in their way. We were particularly interested in the following questions: the following questions: does the system increase the accuracy of user recall? How does the quality of the suggestions affect a user's decision making?

Prior work on context aware memory aids [2,4] indicates that the effectiveness of such systems can be studied in a lab setting, although in neither of these papers are the memory aids related to energy consumption or remembering everyday activities. One key issue is ensuring that the participant is in fact using his or her long term memory and not their short term memory; a solution is to use a distraction activity such as a simple game.

Protocol

Starting from genuine user data and annotations collected in the previous FigureEnergy field trial, we built 4 *narratives*: fictional accounts of day-to-day activities written as third person accounts. To assess how the accuracy of the pattern recognition system affects users' behaviour in a realistic fashion, we selected segments of the original logs for which the recognition accuracy ranged from 17% to 75%.

Each participant was asked to read and memorize one narrative, and then engaged in a distraction task: play a simple computer game (e.g. Minesweeper) for 3 minutes. After the distraction task, the participant is asked to annotate the energy consumption log corresponding to the

narrative they memorised. This sequence is repeated 4 times, each with one of the following conditions: *i)* no suggestions; *ii)* event detections, but no suggested labels as predicted by the system; *iii)* event detections, and suggested labels as predicted by the system; *iv)* suggested event detections, and suggested labels corresponding to the original annotations made by the actual users.

The first is the control condition, it should indicate the baseline recall performance. Conditions 2 and 3 aim at testing the different automation components. Condition 4 corresponds to the "ground truth" and it should reveal whether participants trust the system suggestions or simply ignore them. The order of the conditions is fully counterbalanced and each condition is randomly combined with a narrative, so that overall all narrative appear in all conditions. Performance is measured in terms of the percentage of usage events annotated correctly.

Pilot

A pilot study was run with five participants, some of whom were familiar with the FigureEnergy system, but had no prior experience of the semi-automatic annotation extension. We saw that the distraction activity was successful, as participants were unable to remember all of the events without help from the system.

This pilot suggested clear differences between responses with and without the automated help, showing that accurate suggestions aided recall, improving the participants' annotations. Participants correctly annotated the fewest events under the manual condition in all but one case. The results of this pilot indicate that it is worth running the full study.

REFERENCES

1. Costanza, E., Ramchurn, S.D. and Jennings N.R., Understanding Domestic Energy Consumption through Interactive Visualisation: a Field Study. To appear in *proc. Ubicomp 2012*.
2. DeVaul, R.W., Pentland, A. and Corey, V.R., The Memory Glasses: Subliminal vs. Overt Memory Support with Imperfect Information. In *proc. IEEE ISWC'03*.
3. Fischer, C., Feedback on Household Electricity Consumption: A Tool for Saving Energy? In *Energy Efficiency*, vol. 1 no. 1, Springer (2008), pp. 79-104.
4. Kalnikaite, V. and Whittaker, S., Software or Wetware?: Discovering When and Why People Use Prosthetic Memory. In *proc. CHI'07*.
5. Zeifman, M. and Roth, K., Nonintrusive Appliance Load Monitoring: Review and Outlook. In *IEEE Trans. Consum. Electron.*, vol. 57 no. 1, IEEE (2011), pp. 76-84.